Why Maps Are Silent When Texts Can Speak. Detecting Media Differences through Conceptual Modelling

Øyvind Eide

Unit for Digital Documentation, University of Oslo

Abstract. In this article, a series of experiments will be described in which the expressiveness of maps is shown to be different from the one we find for texts. The hypothesis is that types of geographical information exist that can be stored in and read from a text, but that are impossible to express as geographical maps without significant loss of meaning. The article shows how the hypothesis was supported by a series of modelling experiments.

Keywords: conceptual modelling, textual descriptions, cartography

1. Introduction

The aim of this study is to understand the expressiveness of texts in the area of geographical information better, using topographical maps as a yardstick with which to compare texts. What can a text say, and how is it different from what a map can say? The text used as the source for this study is a published volume of border protocols from 1742–1745 (Schnitler 1962). This article is based on my 2012 PhD from King's College London, Department of Digital Humanities.

The creator of the protocols, Peter Schnitler, initiated a number of court meetings documented in his protocols. Schnitler did not only write the text, he also created the situation in which many of the events described in the text took place. The text documents situations of ordinary, skilled communication about geography, in line with what Mark, Smith et al. (1999) describes. According to the vocabulary developed in the field of regional geography, Schnitler's role was that of a gatekeeper, mediating among a number of different folk geographies existing at the same time:

> [W]ithin any region, there are multiple geographies, varieties of geographical knowledges. ... And there are varieties of folk geographies – those inspired by diverse belief systems and ethnic traditions, as well as those which have emerged from diverse *genres de vie* farmers, fishermen, nomads, merchants and poets (Buttimer, Brunn et al. 1999: 130; italics and missing punctuation in original)

Schnitler also made his own observations during his travels, adding to the experience he already had. In this way, he became an inner voice along with the witnesses as well as the outer voice of professional knowledge. It may be that this mix between being an inner and an outer voice made him more respectful towards the witnesses than a reporter working only from the outside would have been.

Because of the complexity of the historical and cultural background, Schnitler (1962) became a very rich textual expression of detailed geographical information. It is highly descriptive, and includes a number of different perspectives, covering various professions and ethnic groups. As such, it is a good object of study for my experiments.

2. Method

The method applied in this research is experimental modelling as it is used in digital humanities. Models are representations of something which are created for the purpose of studying what is modelled more closely (McCarty 2005). I studied the expressiveness of the source text by manipulating a model of it. Most of the experiments were performed in a qualitative manner, in line with traditional research in the humanities in which knowledge about specific instances of a textual phenomenon is in the first instance the main goal (Galey 2010). Counting and comparing occurrences in the model to find results were done in only a few cases, and the quantitative methods I did use were quite simple.

The modelling was based on the pretence that I knew nothing about the landscape described in Schnitler (1962), except for the descriptions in the text itself. I pretended not to have seen any map and not to know the area, or even general features of Scandinavian landscapes. I pretended not to see most of the context in order to get at what could be read out of the text alone. Avoiding context is impossible in a strict sense, but I limited the context taken into consideration as much as possible.

The process of modelling consists of extracting readings of single assertions made in the text. When a number of such assertions are formalised and interlinked with one another, a surprisingly complex structure is estab-

lished, showing how much there is to even quite commonplace processes and assertions. This also shows how incomplete our knowledge is, and how easily we skip over details in order to get to the big picture. The computer application developed for the modelling, **GeoModelText**[1] offers rich manipulatory power over the interlinked set of assertions. This power has been used to pinpoint details in how the text works and how it is different from topographical maps.

The choice of the verbalised form, that is, 'modelling' rather than 'model', is no accident. The computational model as a fixed structure of knowledge was not a goal in itself, but rather a series of temporary states in a process of coming to know. The point of a modelling exercise lies in the process, not in the model as a product. The use of a computer sharpens the distinction between a model and a concept. The model invites us to manipulate it, as it is set up as an interactive system. Interaction with the computer is doubled when computer programming is a part of the research. In addition to the interactivity one has as a user of a system, one also has the interaction of a developer with the computer.

GeoModelText is not meant to be used for a full analysis and modelling of all information, or even all geographical information, in the source text. It has been created in order to analyse enough to make inferences about how the source text works in the context of the hypothesis.

2.1. Overview of the modelling stages

The modelling process consisted of five main stages. The first stage is the text, imported as a digital document. The next stage is the primary model, consisting of a set of statements in a form inspired by ontologies as they are used in digital cultural heritage, e.g., CIDOC-CRM (CIDOC 2011).

Sentences **I** and **II** are examples of such statements in the form of triples. **I** expresses the fact that the text claims that north of the place referred to by the name "Røvola" there is a place referred to by the referring string "a valley". Sentence **II** expresses the textual claim that the valley from **I** is approximately one-quarter mile wide.

(I) Schn1_8936 (Røvola) → direction: north → node48 (a valley)

(II) node48 (a valley) → has-width → node49 (some ¼ mile)

---

[1] URL: http://geomodel.sourceforge.net/ (checked 2012-11-15)

Based on the primary model, the formalised model is developed. The process from primary to formalised model consists of bringing all the statements to a similar level of explicitness. The next stage is the vector data, where each of the places referenced above, the ones represented by "Schn1_8936" and "node48", are placed in a mathematical vector space based on the relationships between them. Such vector data can be expressed as topographical maps, which is the final stage.

2.2. Stepwise formalisation and fall-off

A system of stepwise formalisation is central to the modelling method I use, and the computer tool **GeoModelText** is implemented to support it. I will show a simple example in order to give an overview of the process. It starts with the short sentence in the "text" column of table 1 and ends with the illustration of map visualisation in figure 1.

Text	Primary model	Formalised model	Vector	Map
Some ¼ mile east of A is B	Unknown mile: 0.25 Direction: east	2 kilometres Direction: 90°	A = (0; 0) B = (2000; 0)	Figure 1

Table 1. Example of stepwise formalisation from text to map.

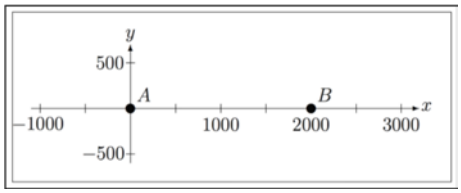


Figure 1. Graphical representation of the model in table 1, showing how the points would be put on a map. Scale in meters.

Different types of expressions will vary in how much they are changed from one stage to the next in the formalisation process. In some cases there is not much added in formalisation from text to primary model. "East" is an example of this. The directions are kept as words in the primary model, formalised into English expressions following a system of 16 directions: north, north-north-east, north-east, etc. Other types of expressions need more formalisation at this stage, such as distance. The number is reformatted from a fraction to a decimal number, and the type of mile is expressed explicitly. The fact that it is not clear from the textual context what type of mile we have is expressed in the name used in the primary model, "unknown mile".

In proceeding from primary to formalised model, a recalculation of the distance from miles to kilometres is done. In the eighteenth century, an official mile in Norway was 11.3 km. Miles used by common people were in general shorter, down to the half of the official, so one unknown mile can reasonably be seen as 8 kilometres. As for the direction, the word "east" is translat-

ed into a number, in this case, 90°. In order to reach the stage of the formalised model, I had to make important choices, for the distance as well as the direction; they could both have been chosen otherwise. This is a key point to which I will return below.

Moving on from the formalised model to vector data, the distance and the direction are taken together. We choose to put A at the origin of the vector room. Once that is done, and as B is 90° from A, and 2 kilometres away, the coordinate for B gives itself. Figure 1 shows the vector data as they would have been expressed on a topographical map.

Information from real chunks of text "travels" through this process, and the method documents what is lost on the way, as well as what must be added in order to reach the goal. Some types of information are lost from one stage to the next because it is impossible to include that information in the version of the model at the next stage. It "falls off". I will use the phrase fall-off to refer to this, both as a verb (to fall off) and as a noun (a fall-off).

Through this process I translate a set of geographical information from one medium to the other. The series of fall-offs shows me what is lost in the process. The process is necessarily iterative and includes a close human-machine interaction. For each small step, the model becomes a little bit more formal. In the process, the fall-off—what is difficult to avoid losing—is the interesting parts. The fall-off will include things that cannot survive a transfer from one stage to the next, which misses the different level of formality. This includes, for example, a direction such as "east". In a system where one must add a specific number of degrees, such as 90, the openness of the word "east" falls off.

Modelling in this project is transitive, so the map is a model of the text. In each step of the stepwise formalisation, the next stage can be seen as a model of each of the previous stages, so that the map is a model of the text.

However, we must remember that even if a map is a model of the text, it is far from the only one possible. There are always other choices that can be made, from the simple level of the choice of a value at a specific step all the way to the overall methodological choices. Even if something falls off in the modelling process from text to map, it does not follow that it cannot be expressed as an element of a map. It could be that it cannot make it through my system of formal models even if it could have been expressed in a topographical map. So the list of fall-offs is not a list of textual features that cannot be expressed in maps, but rather a list of candidates for further study of types of expression that cannot make it to the map.

The automatic vector data generation of **GeoModelText** was not developed to a level where vector datasets, and thus maps, could be created for

anything more than a handful of statements. Thus, the map creation system was only used to demonstrate principles at a simple level, not to document how maps could be made for stretches of text such as a full interview.

3. Case studies

In this section, case studies that demonstrate some of the results found in my PhD research will be presented.

3.1. Witness statements

Povel Olsen's witness statement (Schnitler 1962: 141–143) amounts to 13 paragraphs with a total of 771 words. It includes some 80 references to places, referring to 47 different places.

Looking at the form of the places shown in table 2 first, it is clear that the level of specification is very low. Size is given for only two of the 47 places, and in those two cases only **length**, not **width**, is given. The general **direction** (e.g., that a lake stretches from northeast to southwest) is also given for two places. There is almost no information in the text as to what the places look like, and thus, few clues telling us how they should be drawn. Some size restrictions are given by the relationships to other places, but only in a very vague way. In some cases, other indications related to both the form of a place and the relationships to other places are given, such as rivers connecting to lakes at specified places. But only a few such indications are given, and they are quite vague.

Places		
Type	**Count**	**Relative**
Total number	47	100.0 %
Has length	2	4.3 %
Has width	0	0.0 %
Has direction	2	4.3 %

Table 2. Statistics for form of places.

Relationships between places		
Type	**Count**	**Relative**
Total number	48	100.0 %
With direction	12	25.0 %
With distance	13	27.1 %
Part of	17	35.4 %
Between	14	29.2 %

Table 3. Statistics for relationships between places. The sum of the different types of relationships is higher than the total number of relationships because many of the relationships have both direction and distance specified.

Most of the 47 places do have spatial connections to other places, but the spatial specificity varies. Table 3 shows the different connection types and the number of connections for which each of them are specified. When both **direction** and **distance** are given, we have the most specified relationships. **Part of** and **between** are less specific, spatially speaking. But these are differences in grade; all the relationships

between places are underspecified. An example showing how underspecification works can be seen in figures 2 and 3, where two different maps based on the same textual description, taken from (Schnitler 1962: 141), are shown. The different rules of interpretation on which they are based are:

1. South is interpreted as 180° (fig. 2) or 160° (fig. 3).

2. Default width and length of a place polygon when no measurements are given: 4000*6000 (fig. 2) or 1000*500 (fig. 3) meters.

3. Default distance in *X* and *Y* direction of something between something else, when distance and direction are not given: 500*200 (fig. 2) or 1000*2000 (fig. 3) meters.

4. The length of a mile when type is not specified: 8000 (fig. 2) or 6000 (fig. 3) meters.

Figure 2. Map based on textual description, version 1. Scale in metres.

Some non-spatial relationships can also be put on a topographical map; e.g., if two places are related only by both being farms, the relationship can be expressed by the use of the same symbol on the map. However, if there is no spatial relationship indicating where one is located relative to the other, the places cannot be put on the same map in any other way than by guessing the location of one of them relative to the other. The fact that both of them are farms tells us that they can be represented by the same symbol, but not their spatial relationship.

Figure 3. Map based on textual description, version 2. Scale in metres.

In a communication situation it is common to mention a well-known place without specifying where it is in relation to other places. This will still give its location to an informed reader. A place name should, when it is used within the group of people knowing it, "instantly evoke ...the idea of one particular place through an association by c o n t i g u i t y" (Olsen 1928: 5; highlighted in original).

The place names in Schnitler (1962) give the locations of the places through such associations, and through the locations thus given, spatial relations to other places are also implied. These relations are not found in the text, but in the knowledge of the reader. The writer is also a reader under this perspective: he is the first reader. Some of the readers, even the writer, may not

know where some of the places are. Then they will have to fall back on the spatial relations given by the text in order to get an idea of the whereabouts of the places based on other places whose location they do know. In my modelling, a situation where the reader knows about none of the places referred to by the place names is simulated. There is no way to choose one and only one correct way of expressing what we are able to put on a map based on the text only.

This overall pattern is repeated throughout the text. There are, however, some room for variation. Another witness, Ole Nilsen, shows somewhat higher levels of connectedness. In his interview, there are a total of 215 references to places (Schnitler 1962: 150–155). These 215 expressions refer to 115 different places. The total number of relationships between places is 155. The specificity of the relationships is higher than the one we found for Povel. The percentages of distance and part of are more or less the same, but there are more directions at the expense of between here, which gives a higher level of specificity.

When we look at descriptions of places, the difference is even clearer. Length, width and direction are all used for more than 10% of the places in Ole's testimony, compared to less than 5% for Povel's. So Ole describes form more than Povel. Still, even Ole gives no such descriptions of form for almost 90% of the places. This indicates that even if there is room for some variation in descriptions of the form of places, the level is still low. The witnesses depend on the instant evocation of the particular place in the mind of the reader—perhaps not consciously, but at least such dependence is part of their way of expressing themselves.

3.2. Route descriptions

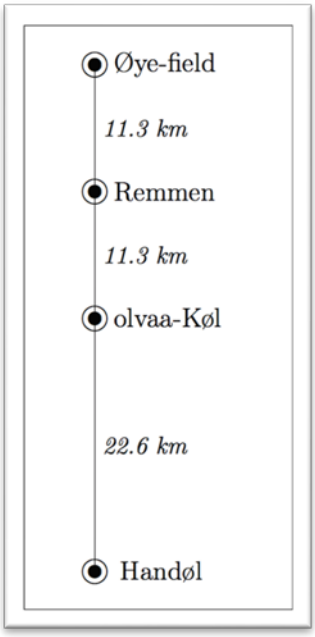
The road to *Jemteland* the *ordinary one* goes from *Østbye* Farms to Øye mountain can be estimated	1 new mile,
from there through a little birch forest by the western end passing by the Lake *Eesand* in the east to *Remmen*	1 mile
from there through a little birch forest to *olvaa-Køl*	1 mile
from here through the forest to *Handøl* the 2 first farms in *Jemteland*	2 miles
---	
New	5 miles —

Figure 4. English translation of a route description (Schnitler 1962: 54).

Route descriptions are common in the text. In the route description from Norway to Sweden shown in figure 4, the directions of the roads are not given in the text. If the reader takes the wider context into consideration, the general direction would have been known, because Sweden is generally east of Norway in this area. But this is not expressed in the text fragment. In line with what we saw above we cannot make a topographical map representing the route based on this information alone; several choices have to be made. We would, for instance, have to choose an exact direction for each of the relations.

Instead of trying to express the model of this text fragment as a topographical map, I tried a topological map instead. The text of the route description does not contain enough information to produce one specific topographical map, but it turned out that one and only one topology is readable from it; thus, two significantly different topological maps cannot be made, as long as we see "significantly different" as "expressing different topologies".

The text does not as much describe a landscape, which would be translatable to a topographical map, as a network, which is translatable to a topological map. In figure 5, such a topological map is shown. In addition to what we usually have in



Figur 5. A topological map based on some of the text in figure 4. Topological maps are usually not scaled, but this one is.

topological maps, the text also informs us about the relative distances, so that the topological map in this case can be made to scale; the road between one pair of places is double the length of the two other roads. But there is a difference between the topological connections between the places, on one hand, and the scale, on the other. The former is precisely as it was described in the text, whereas the latter has the inaccuracy inherent in all distances in the text, and thus, the interpretation can be chosen to be different.

Route sketch maps have previously been shown to be based on the same mental models as textual route descriptions (Tversky and Lee 1999). That confirms the similarity between a textual description and a topological map shown here. A topological map is a good way to express a network with only some spatial constraints; the same can be said of a route description.

4. Conclusion

When creating maps based on a text using previously existing maps and local knowledge, it is often doable to express information from the text unambiguously on the map. But in that case one locks the text to one geographical reading. This may be problematic for historical texts. While I do not propose to use the method described here in normal historical text studies, I would suggest the results are taken as a further evidence of how much of what we read is based on the knowledge with which we meet the text. Many choices must be made, also when the choices are linked to a choice of which map to use.

For most spatial features, there is more or less leeway in how they can be interpreted if we rely on the text alone. A statement giving only a very limited leeway would be “A is exactly 1.03 km from B, in a direction of exactly 87.6°”. There are no such statements in our source text, the relationships are significantly less specified if expressed at all. However, the semiotic system of a topographical map calls for a high level of precision, which must be added for the map to be created.

Although there are ways to at least partially tackle the problems described in this paper, such as fuzzy logic (Petry, Cobb et al. 2005), the differences between texts and maps still remain. Fuzzy forms on maps are commonly used for thematic mapping, but less so in historical mapping of topographical features other than for inherently fuzzy objects such as mountains. It may take the development of new map schemata (MacEachren 2004: 198) to change this. And the question remains to what degree new ways of understanding maps can bridge the gap between the semiotic systems of texts and maps.

The results described here call for further research. In order to understand textual underspecification better, more texts from a variety of cultures should be studied, in light of previous textual studies (Eide Forthcoming 2013). Further, better methods for modelling should be implemented in order to deepen our understanding of what can be expressed in the different media. Dynamic maps should also be taken into consideration.

References

Buttimer, A., et al. (1999). III. Polyphonies of Place and Indetity: Mapping Self and Other. Introduction. Text and image: social construction of regional knowledges. A. Buttimer, S. D. Brunn and U. Wardenga. Leipzig, Institut für Länderkunde**:** 130--2.

CIDOC (2011). Definition of the CIDOC Conceptual Reference Model. [Heraklion], CIDOC.

Eide, Ø. (Forthcoming 2013). Verbal expressions of geographical information. New Worlds out of Old Texts: Developing Techniques for the Spatial Analysis of Ancient Narratives. E. Barker, S. Buzar, C. Pelling and L. Isaksen. Oxford, Oxford University Press.

Galey, A. (2010). The Human Presence in Digital Artefacts. Text and Genre in Reconstruction: Effects of Digitalization on Ideas, Behaviours, Products and Institutions. W. McCarty, Open Book Publishers**:** 93-117.

MacEachren, A. M. (2004). How maps work : representation, visualization, and design. New York, Guilford Press.

Mark, D. M., et al. (1999). Ontology and Geographic Objects: An Empirical Study of Cognitive Categorization. Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. C. Freksa and D. M. Mark. Berlin, Heidelberg, Springer-Verlag**:** 283--98.

McCarty, W. (2005). Humanities computing. Basingstoke, Palgrave Macmillan.

Olsen, M. (1928). Farms and fanes of ancient Norway: the place-names of a country discussed in their bearings on social and religious history. Oslo, Aschehoug.

Petry, F. E., et al. (2005). Fuzzy Modeling with Spatial Information for Geographic Problems. Berlin, Heidelberg, Springer-Verlag Berlin Heidelberg.

Schnitler, P. (1962). Major Peter Schnitlers grenseeksaminasjonsprotokoller 1742--1745. 1. Oslo, Norsk historisk kjeldeskrift-institutt.

Tversky, B. and P. U. Lee (1999). Pictorial and Verbal Tools for Conveying Routes. Spatial information theory : cognitive and computational foundations of geographical information science : International Conference COSIT '99, Stade, Germany, August 25-29, 1999 : proceedings. C. Freksa and D. M. Mark. Berlin, Springer-Verlag: 51--64.